

Using Embeddings to Compare FrameNet Frames Across Languages

Jennifer Sikos

IMS, University of Stuttgart
Stuttgart, Germany

sikos@ims.uni-stuttgart.de

Sebastian Padó

IMS, University of Stuttgart
Stuttgart, Germany

pado@ims.uni-stuttgart.de

Abstract

Much of the recent interest in Frame Semantics is fueled by the substantial extent of its applicability across languages. At the same time, lexicographic studies have found that the applicability of individual frames can be diminished by cross-lingual divergences regarding polysemy, syntactic valency, and lexicalization. Due to the large effort involved in manual investigations, there are so far no broad-coverage resources with “problematic” frames for any language pair.

Our study investigates to what extent multilingual vector representations of frames learned from manually annotated corpora can address this need by serving as a wide coverage source for such divergences. We present a case study for the language pair English — German using the FrameNet and SALSA corpora and find that inferences can be made about cross-lingual frame applicability using a vector space model.

1 Introduction

Frame Semantics (Fillmore et al., 2003) is a theory of predicate-argument structure that describes meaning not at the level of individual words, but is instead based on the concept of a scenario or scene called a *frame*. Frames are defined by the group of words that evoke the scene (*frame-evoking elements* or *FEEs*), as well as their expected semantic arguments (*frame elements*). A JUDGMENT frame, for instance, would have FEEs *praise*, *criticize*, and *boo* and frame elements such as COGNIZER, EVALUEE, EXPRESSOR and REASON. The Berkeley FrameNet project (Baker et al., 1998) is the most well-known lexical resource based on Frame Semantics, and provides definitions for over 1200 frames.

A very attractive feature of Frame Semantics is that it can abstract away from individual words, which makes it a promising descriptive framework for events that generalize across languages (Boas, 2005; Lönneker-Rodman, 2007). However, there is a fundamental tension within Frame Semantics between the characterization of the scene, which is arguably relatively language-independent, and the characterization of the FEEs and frame elements which can vary among even closely related languages. For example, FrameNet distinguishes between the frames OPERATE_VEHICLE (*drive*) and USE_VEHICLE (*ride*) which are often indistinguishable in German, at least for the very frequent verb *fahren* (*drive/ride*) (Burchardt et al., 2009). Another example is the well-known difference between Germanic and Romance languages regarding the conceptualization of motion events, where manner can be incorporated into the semantics of the FEE - *run into a room*, or externalized - *enter a room running* (Subirats, 2009). In general, some frames are evidently more parallel than others cross-lingually (Tonelli and Pianta, 2008; Torrent et al., 2018; Vossen et al., 2018).

The question of whether a particular frame transfers across languages is often determined by linguists on the basis of large corpora, comparing either frames assigned to direct translations (Padó, 2007) or extracting “typical” uses of frames in the languages under consideration. Identifying the degree of applicability can, however, be a long process that must be repeated for each new frame and language.

In this paper, we investigate to what extent this problem can be alleviated by the use of computational methods. More concretely, we build distributional representations, also known as embeddings, of frames

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>

(Turney and Pantel, 2010; Baroni and Lenci, 2010; Hermann et al., 2014) which build a compact representation of a frame’s meaning from the entire corpus. These embeddings are constructed for each language separately on the basis of expert-labeled corpora but are represented in a joint space, which enables us to automatically measure the cross-lingual similarity of language-specific frames (e.g., AWARENESS in English and AWARENESS in German). Our hypothesis is that frames which generalize well across languages will show a high similarity, while frames which do not generalize well will yield a low similarity. In this sense, we can see the set of frames that show the lowest self-similarity across languages as a repository of cases of potential generalization failure. Concretely, we carry out a case study for the language pair English–German, which is relatively well-studied in terms of Frame Semantics.

This paper is organized as follows. First, we describe distributional semantic models (*DSMs*) and their utility in capturing cross-lingual lexical semantics. We give an overview of the annotated corpora that we use for our experiments for English – German frames, followed by an overview of the model structure and input for training our DSMs. Finally, we present results of our frame embeddings in both English and German data and interpret output of our computational model through findings from the theoretical literature. To our knowledge, our work is the first to explore interoperability of frames across languages with embeddings. We present results that demonstrate how DSMs can capture latent semantic relationships between frames and speculate about the conclusions that can be drawn when using DSMs for frame semantic analysis.

2 Background

2.1 Word Embeddings

Vector representations of word meanings, generally known as distributional semantic models (*DSMs*) or embeddings, have become a popular tool for research in lexical semantics within the Natural Language Processing community. The distributional hypothesis (Harris, 1954) forms the basis for DSM models, which claims that similar words should be expected to appear in similar contexts. Building on this assumption, current DSMs represent words as vectors in a low-dimensional space in which similarity is supposed to indicate semantic relatedness. DSMs can either be computed using count-based methods, which directly count word context co-occurrences and optionally apply dimensionality reduction, or prediction-based methods, where the representations are the parameter vectors of an underlying optimization task (Baroni et al., 2014).

A widely-used prediction-based model is the word2vec model (Mikolov et al., 2013c). Word2vec produces word vectors by predicting the target for a window of context, a setup known as the continuous bag-of-words (CBOW) model. In the CBOW architecture, an input for the target verb “cook” with a context window of 2 might appear as [‘the’, ‘chef’, ‘a’, ‘meal’], and embeddings would accordingly be generated for each token in the vocabulary. The model is trained over a neural network with negative sampling, where noisy words are selected from a noise distribution and the model learns to approximate word meaning by discriminating the difference between the true and noisy word.

While the use of embeddings to model meaning at the word level is the most widespread application, they can in fact be applied to linguistic units of almost any granularity, from morphemes to word senses to sentences and even paragraphs and documents. In this study, we apply them to modeling the meaning of frames, i.e., semantically motivated predicate classes, following Hermann et al. (2014).

2.2 Cross-lingual Embeddings

The goal of a cross-lingual DSM is to make meaning representations from different languages comparable, either by learning multilingual embeddings in the same space or by mapping monolingual embeddings into a joint space (Ruder et al., 2017; Upadhyay et al., 2016). Mikolov et al. (2013b) introduced a particularly simple version that takes advantage of a vocabulary of shared bilingual seed words to map embeddings from a source language onto the vector space of a target language. They learn a transformation matrix \mathbf{W} by minimizing the mean squared error (MSE) between the source seed words and their translations from the bilingual dictionary:

$$\text{MSE} = \sum_{i=1}^n \| \mathbf{W}x_i^s - x_i^t \|^2$$

where x^s represents the seed words from the source language, and x^t are seed words from the target. The transformation matrix can then be applied to an embedding from the source space to project it onto the target language space. This can be achieved by solving $z = \mathbf{W}x$, where x is the source embedding and z is the same vector transformed into the target language space. This results in a shared space where semantic units that evoke similar concepts in one language are close to their translations in another.

Conceptually, the bilingual seed words is a set of words that are assumed, a priori, to have the same meaning in both languages. A good choice for this set appear to be named entities, which are less prone to semantic problems like polysemy and vagueness, and which are furthermore easy to collect for most language pairs. We will adopt this choice in the present study.

3 Annotated Corpora

The popularity of the Berkeley FrameNet project has inspired the Multilingual FrameNet project, where frame-semantic resources have emerged in multiple languages across the world (Torrent et al., 2018). In our work, we focus on two languages that are relatively well-represented in the Frame Semantics literature, English and German, using the FrameNet and SALSA corpora, respectively.

3.1 FrameNet Corpus

The FrameNet corpus (release 1.5) provides over 173k annotated datasets with manually-curated sentences, drawn from the 100 million word British National Corpus (BNC)¹, which is balanced across multiple genres, including novels and news reports. FrameNet’s FEEs can range from verbal, nominal, adjectival, and even adverbial and prepositional in some cases. In FrameNet’s v1.5 complete annotation set, there are over 11k FEEs, where each FEE attestation is labeled. This produced sentences in which multiple FEEs are present, and lexical units are annotated on average 20.7 times in the current annotated corpus.

3.2 SALSA Corpus

SALSA (Erk et al., 2003; Burchardt et al., 2009) is a German corpus with frame semantic annotations. It was constructed over the syntactic annotations in the German TIGER (version 2.1), and is composed of 1.5 million words from German newswire text. SALSA (release 2.0) provides 24,184 sentences with frame annotations, and the corpus has labels for 998 unique FEEs. Each unique FEE is annotated on average 36.9 times.

3.3 Parallelism

For the purposes of our study, it is important that the corpora are as comparable as possible. Comparability must be considered on two levels: corpus composition and annotation. With regard to corpus composition, we note that SALSA is based on a corpus consisting exclusively of newswire, while FrameNet is based on the much broader BNC. Nevertheless, we maintain that the match is good, albeit not perfect. As for the semantic annotation schema, SALSA adopted the annotation scheme of the Berkeley FrameNet project fairly directly (Ellsworth et al., 2004). The main difference is that SALSA proceeded predicate-by-predicate instead of frame-by-frame, which required the introduction of “pseudo-frames” for senses that were not covered by FrameNet at the time of annotation (Burchardt et al., 2009). We exclude these from our analysis. Finally, another small difference regards the annotation of multi-word expressions, which we address further in Section 4.

4 Methods

To build frame embeddings for English and German, we pre-process the two corpora introduced in the previous section. Below, we describe the steps we took to format the input for the model, and then we elaborate on how specifically we build the frame embeddings.

¹<http://www.natcorp.ox.ac.uk/>

4.1 Preprocessing the FrameNet corpus

We use the FrameNet 1.5 corpus, which provides tokenized and lemmatized sentences (Bauer et al., 2012). We perform three pre-processing steps. First, multi-word frame-evoking elements (FEEs) are concatenated into a single token. This is motivated by the presence of multi-word FEEs that evoke different frames from the single-word FEEs that they contain. One such case is the lexical unit *ride*, which evokes the RIDE_VEHICLE frame, and the multi-word FEE *ride out*, which evokes SURVIVING. We keep instances of MWE FEEs such as *ride out* as single units in the vector space.

The second step is a detection of named entities using the pretrained English model from spaCy’s named entity recognition (NER) package². We later use these entities as seed words for the cross-lingual mapping of embeddings (cf. Section 2). Again, for simplicity, we concatenate entities that span multiple tokens into a single span, e.g., “San Francisco” would be converted to “San_Francisco”.

The third step optionally replaces each occurrence of a FEE by the name of the frame it evokes to produce the “Frame Corpus”. By applying the word2vec embedding algorithm to the Frame Corpus, we can generate a monolingual embedding for each frame. Without this replacement, we retain the so-called “FEE Corpus”, from which we can generate standard word-level representations for the FEEs. An example of the FEE Corpus and Frame Corpus model inputs are shown below in Table 1.

| | |
|--------------|---|
| Original | “The Washington Post reported on the country’s biological weapons labs” |
| FEE Corpus | [The_Washington_Post, report, on, the, country, ’s, biological_weapon, lab] |
| Frame Corpus | [The_Washington_Post, STATEMENT, on, the, country, ’s, WEAPON, lab] |

Table 1: Example of original FrameNet sentence and the FEE Corpus and Frame Corpus versions

4.2 Preprocessing the SALSA corpus

Similar to FrameNet, the SALSA corpus provides pre-tokenized and lemmatized sentences (Erk et al., 2003). Multi-word predicates are prevalent in the SALSA data but cannot be concatenated as easily as they are in English; SALSA decided to annotate some constructions as FEEs that are arguably more syntactically than semantically motivated. Table 3 shows the three most frequent MWE patterns. The first class, separated prefix verbs, are merged into a single word. The second class, combinations of verbal FEEs with modals, and the third class, combinations of verbal FEEs with the particle *zu* (*to*), we decided not to concatenate. Like English, we recognize named entities using spaCy (using the pretrained German model in this case), and convert any named entities that span over multiple words into a single token (“Volksrepublik China” becomes “Volksrepublik_China”). Also identical to the English FrameNet input, we build two corpora using the German data, a FEE Corpus and a Frame Corpus, to generate embeddings for frames and FEEs, respectively. An example is shown in Table 2.

| | |
|--------------|---|
| Original | “Konzernchefs lehnen den Milliardär als US-Präsident ab” (CEOs reject the billionaire as US President) |
| FEE Corpus | [Konzernchef, ablehnen, der, Milliardär, als, US-Präsident] |
| Frame Corpus | [Konzernchef, JUDGMENT_COMMUNICATION, der, Milliardär, als, US-Präsident] |

Table 2: Example of original SALSA sentence and the FEE Corpus and Frame Corpus versions

4.3 Frame Embeddings

For both English and German, we first learn 300-dimensional monolingual FEE-level and frame-level embeddings from the respective corpora, using the word2vec CBOW method (see (Mikolov et al., 2013a) for more in-depth details). Then, we employ Mikolov (2013a)’s method based on a shared bilingual vocabulary to learn a linear projection from the source to the target vector space. As a bilingual vocabulary, we use the intersection of the recognized named entities of the standard classes (organizations; persons

²<https://spacy.io/>

| | | |
|-------------------------|---|-----------|
| VVFIN+PTKVZ | → | PTKVZ_VV |
| Example: gehört an | → | angehören |
| VM* VV* | → | no change |
| Example: müssen rechnen | | |
| PTKVZU | → | no change |
| Example: zu sagen | | |

Table 3: Rules for combining multi-word predicates in German SALSA corpus

| | |
|------------------|----------------|
| Bill Clinton | Mussolini |
| Netanyahu | Vegas |
| Pentagon | CIA |
| Intel | IBM |
| San Francisco | McDonald’s |
| Apple | Bill Gates |
| Mao Zedong | Reuters |
| Washington Times | New York Times |

Table 4: Examples of named entities occurring in both FrameNet and SALSA

and person groups; locations and geopolitical entities) from FrameNet and SALSA. In other words, we first match on named entities whose surface form is identical between English and German – so *New York / New York*. To increase the size of the seed entities, we use a bilingual dictionary for English–German (Conneau et al., 2017) to further match any named entities where the surface form varies but reflects the same individual or location, for example *President Clinton / Präsident Clinton*. Overall, 40,262 entities were detected in the English FrameNet sentences, while 15,398 entities were detected in the German SALSA. The intersection of entities that appeared in both corpora was 2,899. The list in Table 4 gives a sample of the entity names that appear in both FrameNet and SALSA.

The models are trained on the pre-processed corpora. This results in a total of 230 frame embeddings for both EN and DE, 11,830 English FEE embeddings, and 998 German FEE embeddings.³

5 Experiment 1: Hyperparameter optimization and sanity check

Producing embeddings with word2vec’s CBOV algorithm involves choices regarding several hyperparameters, notably the number of negative samples (*neg*), the learning rate (*alpha*), the size of the context window (*win*), and the number of iterations (*iter*) – cf. Section 2. We use the same hyperparameter values for the Frame Corpus-based and FEE Corpus-based embeddings.

It is common practice to choose the values for these hyperparameters based either on performance on a development set or on some auxiliary task, such as word similarity prediction. In order to validate the FrameNet/SALSA frames that we use in our subsequent analysis, we decided to define an auxiliary task; namely, a check for the quality of the monolingual frame embeddings. An added benefit of this auxiliary task is that it can also be understood as a form of sanity checking for the monolingual embeddings before using them for cross-lingual comparison.

More concretely, we ask how well the embedding we obtain for each frame corresponds to the centroid of all FEE embeddings of that frame. For example, we would expect that the `COMMERCE_BUY` frame embedding should be highly similar to the centroid of the embeddings of the FEEs *buy*, *purchase*, but dissimilar to centroid of embeddings for predicates such as *tell* or *beat*. Formally, let $c_{FEE}(\phi)$ denote the *FEE centroid* for a frame ϕ , defined as

$$c_{FEE}(\phi) = \frac{1}{|\{f_{ee} | f_{ee} \in \phi\}|} \sum_{f_{ee} \in \phi} \vec{f}_{ee}.$$

We can now assemble, for each frame f , the list of FEE centroids c_{FEE} of all frames ranked by their cosine similarity to the frame embedding \vec{f} . This list enables us to define the accuracy of the model using standard evaluation measures for ranked retrieval. We report the percentage of frames whose most similar FEE centroid is their own (R@1), the percentage whose FEE centroid is among the five most similar ones (R@5), and the percentage whose FEE centroid is among the ten most similar ones (R@10).

5.1 Results

Tables 5 and 6 give a sampling of the some of the hyperparameters that were used to evaluate the frame embeddings. The numbers indicate that the choice of hyperparameters is in fact important. Furthermore, the behavior of the embeddings with respect to the hyperparameters is relatively parallel between English and German: too few iterations and too few negative samples result in noisy embeddings. Setting *neg* to

³The embeddings are available at <http://www.ims.uni-stuttgart.de/data/XLFrameEmbed.html>.

| neg | alpha | win | iter | R@1 | R@5 | R@10 |
|-----|-------|-----|------|--------------|-------|-------|
| 5 | .025 | 2 | 10 | 0.481 | 0.701 | 0.766 |
| 10 | .025 | 5 | 20 | 0.832 | 0.951 | 0.970 |
| 10 | .05 | 2 | 20 | 0.733 | 0.923 | 0.957 |
| 20 | .025 | 2 | 30 | 0.931 | 0.987 | 0.993 |
| 20 | .025 | 2 | 35 | 0.929 | 0.990 | 0.994 |
| 30 | .025 | 2 | 30 | 0.912 | 0.988 | 0.993 |

Table 5: FrameNet (EN): Evaluation of Frame Embedding Quality

| neg | alpha | win | iter | R@1 | R@5 | R@10 |
|-----|-------|-----|------|--------------|-------|-------|
| 5 | .025 | 2 | 10 | 0.408 | 0.591 | 0.653 |
| 10 | .025 | 2 | 20 | 0.843 | 0.925 | 0.938 |
| 10 | .05 | 2 | 20 | 0.938 | 0.986 | 0.993 |
| 20 | .025 | 5 | 20 | 0.904 | 0.972 | 0.979 |
| 20 | .025 | 2 | 35 | 0.965 | 0.979 | 0.986 |
| 30 | .025 | 2 | 30 | 0.931 | 0.993 | 0.993 |

Table 6: SALSAs (DE): Evaluation of Frame Embedding Quality

| Frame | Top 10 Most Similar Frame Embeddings | |
|--------------|---|--|
| | FrameNet (EN) | SALSAs (DE) |
| COMMERCE_BUY | COMMERCE_SELL, GETTING, SUPPLY, IMPORTING, DEGREE_OF_PROCESSING, TRANSFER, IMPORT_EXPORT, RECEIVING, AMASSING, CAUSE_MOTION | GETTING, IMPORT_EXPORT, RECEIVING, CAUSE_MOTION, REMOVING, MANUFACTURING, ACTIVITY_START, USING, COMMITMENT, BRINGING |
| DECIDING | COMMUNICATION_RESPONSE, ASSESSING, SOLE_INSTANCE, ACTIVITY_ONGOING, ATTEMPT, TELLING, GIVING, ARRIVING, DESIRING | COMMUNICATION_RESPONSE, ASSESSING, ACTIVITY_ONGOING, ATTEMPT, TELLING, GIVING, ARRIVING, DESIRING, RESPONSE, PERCEPTION_ACTIVE |

Table 7: Top 10 nearest neighbors for FrameNet/SALSAs frame embeddings

20 and running for 30 to 35 iterations, however, yields arguably good embeddings with R@1 values of over 90%, that is, the vast majority of frame embeddings are located closest to their own FEE centroid. For R@10, the numbers approach (for German) or even exceed (for English) 99%. Note that these tables include all frames that appear in the corpora. When we prune our scoring to only frames that occur >5 times in the corpus, we obtain R@1 scores of over 97% for both languages. For the remainder of the paper, we adopt the hyperparameters that yield the highest accuracy.

We complement these results with a more qualitative analysis, where we find the top ten nearest neighbor frame embeddings for two frames in the same, monolingual space. This determines to what extent the frame embeddings form sensible semantic neighborhoods: We would expect a frame like COMMERCE_BUY to have COMMERCE_SELL as one of its nearest neighboring frames, while a frame like GIVING_BIRTH would not. The results are shown in Table 7 and demonstrate that essentially all nearest neighbors are semantically related to the target frame, expressing a wide range of concepts. For example, for COMMERCE_BUY, we find the inverse perspectivization COMMERCE_SELL, entailments (GETTING, RECEIVING, CAUSE_MOTION), preconditions (COMMITMENT), scenarios and narrative chains (IMPORTING, IMPORT_EXPORT, MANUFACTURING), and aspectual class (ACTIVITY_START). At the same time, the lists are sufficiently dissimilar to motivate our second experiment which assesses to which extent these differences reflect cross-lingual differences in frame usage and applicability.

6 Experiment 2: Analysis of Cross-Lingual Frame Correspondences

As described in Section 4, we project the FrameNet and SALSAs frame embeddings into the same space. This means we can compare them via a simple distance metric such as cosine similarity, and we can visualize them via dimensionality reduction to two dimensions. We start our analysis by confirming that the joint space respects the monolingual similarities and introduces reasonable cross-lingual ones. For clarity, we will affix -EN to English (FrameNet) frames and -DE to German (SALSAs) frames.

Figure 1 shows the nearest frame neighbors for VERDICT in English and a joint FrameNet-SALSAs space. The joint space, while introducing related frames in the German space, still preserves the relationships on the source side; namely, ARREST, NOTIFICATION_OF_CHARGES, and TRIAL. Additional related frames emerge from the German data, including CRIMINAL_INVESTIGATION and PROCESS_START.

We now proceed to analyze our embeddings through the perspective of prior cross-lingual studies

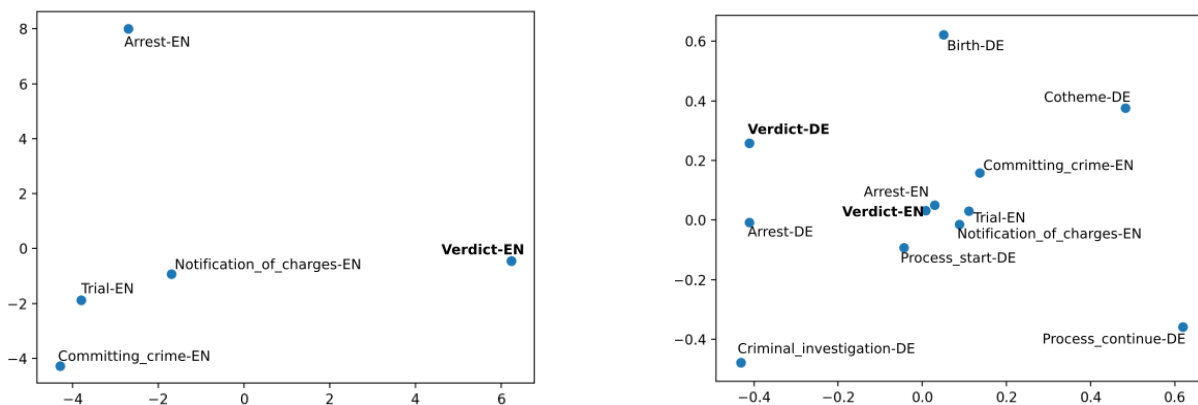


Figure 1: English space around VERDICT-EN (left), and joint space (right). Many semantic relationships are preserved in the cross-lingual mapping.

from the linguistics literature. We focus in particular on the frames and frame properties that diverge in English and German, and compare these findings with results from our corpus-based frame embeddings. Conversely, we explore frames that have a high potential for universality and compare their similarities in the cross-lingual vector space.

6.1 Specificity vs. Cross-lingual Frame Applicability

FrameNet frames form a hierarchy with more abstract, schematic, often scenario-like frames at the top and more specific and more constrained frames at the bottom (Fillmore et al., 2003). Arguably, the more specific a frame is, the higher is the risk that its constraints do not carry over well to other languages, while frames that are higher on the FrameNet hierarchy should be more universally applicable (Padó, 2007).

As two examples for such general frames, consider COMMUNICATION and MOTION, which apply to general speaking and moving events (with FEEs such as *speak*, *communicate*, *say* and *move*, *come*, *go*, respectively). Both frames are extended by more specific child frames, such as COMMUNICATION_RESPONSE, COMMUNICATION_MANNER, MOTION_DIRECTIONAL and SELF_MOTION. We now test whether the specificity of the frames correlates with the cross-lingual similarities of the frame embeddings. We expect MOTION-DE and MOTION-EN to be highly similar while MOTION_DIRECTIONAL-DE and -EN are less similar when mapped into a joint, multilingual vector space.

The results are visualized in Figure 2: in general, both coarsely defined frames have a higher similarity than their respective child frames, although this tendency is much more pronounced for COMMUNICATION than MOTION. The one exception to this trend is the SELF_MOTION frame, which can also be argued is a general frame for self-propelled movement; in fact, the SELF_MOTION is annotated more in the SALSA corpus (count=76) with a greater variety of FEEs (31 unique FEEs for SELF_MOTION in SALSA) than MOTION (count=32), which has significantly fewer unique FEE annotations (10 unique FEEs).

6.2 Most and Least Similar Frames

Next, we computed the frames that were most and least similar across FrameNet and SALSA. Our expectation was that the list of most similar frames would be dominated by very general frames (as per the previous analysis) and that the list of least similar frames would contain mostly cases known as problematic from the lexicographic literature.

These expectations were only confirmed to an limited extent. Amongst the very similar frames, frames tend to be belong to abstract domains such as Cognition (GRASP, JUDGMENT, MEMORY) or describing abstract properties of events (ACTIVITY_RESUME, LIKELIHOOD). Two concrete frames are also prominent: VERDICT and ARREST. This was a surprising result, given the differences in the judicial systems of the USA and Germany.

Even for very similar frames, it is striking how little correlation there is between the FrameNet and SALSA frequencies for certain frames, and the imbalance is more pronounced for the least similar frames.

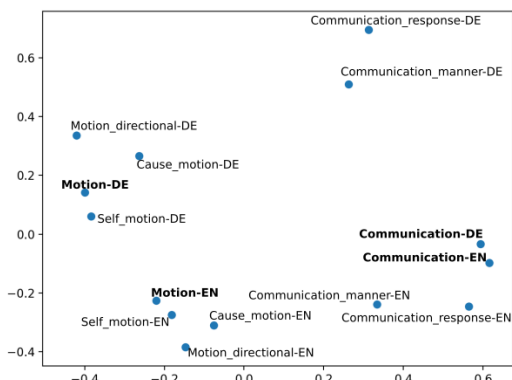


Figure 2: Cross-lingual similarity scores for COMMUNICATION and MOTION.

| Frame | Similarity DE-EN |
|------------------------|------------------|
| COMMUNICATION | 0.54 |
| COMMUNICATION_RESPONSE | 0.17 |
| COMMUNICATION_MANNER | 0.14 |
| MOTION | 0.39 |
| CAUSE_MOTION | 0.15 |
| MOTION_DIRECTIONAL | 0.27 |
| SELF_MOTION | 0.46 |

Figure 3: Cross-lingual similarity between the less-specific COMMUNICATION and MOTION frames and more specific child frames.

| Most Similar Frames | Similarity | Freq EN/DE | Least Similar Frames | Similarity | Freq EN/DE |
|---------------------|------------|------------|----------------------|------------|------------|
| GRASP | 0.66 | 287/21 | ORIGIN | -0.14 | 194/8 |
| COMMUNICATION | 0.54 | 84/25 | PEOPLE_BY_VOCATION | -0.11 | 586/12 |
| VERDICT | 0.51 | 215/77 | UNDERGOING | -0.07 | 38/136 |
| ARREST | 0.51 | 189/103 | INGEST_SUBSTANCE | -0.02 | 187/7 |
| BUILDING | 0.49 | 393/52 | EMPLOYING | 0.00 | 151/428 |
| ACTIVITY_RESUME | 0.49 | 37/8 | TEXT | 0.03 | 1080/4 |
| JUDGMENT | 0.48 | 1212/33 | SENSATION | 0.03 | 471/1 |
| MEMORY | 0.48 | 209/41 | TAKING_SIDES | 0.04 | 189/18 |
| COTHEME | 0.48 | 665/7 | COMMITTING_CRIME | 0.07 | 48/55 |
| LIKELIHOOD | 0.48 | 577/6 | SENTENCING | 0.07 | 77/39 |

Table 8: Most and least similar English–German frame pairs in joint space.

In fact, we believe that for a number of frames in this list (e.g., SENSATION), the reason for low similarity is that SALSA provides so few annotations that no reliable embeddings can be learned, irrespective of any linguistic or lexicographic considerations. For the remaining frames, the low similarity appears to result from large differences in the FEEs that were chosen for annotation. For example, FrameNet’s ORIGIN contains many nationalities and ethnicities (*American, Assyrian, Byzantine*) while SALSA annotated family terms like *Kind, Sohn, Enkel (child, son, grandson)*. Similarly, PEOPLE_BY_VOCATION, which in FrameNet covers a wide range of professions, only contains one FEE in German, *Kumpel (miner)*. These appear to be artifacts of SALSA’s strategy of annotating lemma by lemma, as a result of which some frames end up being annotated only for a single, or very few, FEEs (Ellsworth et al., 2004). So, the least similar frames provide information about annotation choices rather than conceptual applicability.

Note that the definition of abstract/general frames can be made more concrete by considering FrameNet’s frame-to-frame relations “X is Inherited by/is Used by/has Subframe Y”. All relations share the common property that X is a more abstract frame than Y. Thus, if there is a high number of relations in which a frame occurs in position X, then it should be a more abstract frame, and behave more similar across languages. This turns out to be true in our sample: the ten most similar frames listed in Table 8 filled a high average number of relations in position X (4.89) while the ten least similar frames had a comparably low average (1.7).

6.3 Monolingual Annotation Choices

Nevertheless, the previous analysis raises the question of how monolingual annotation choices influence cross-lingual comparability. As an example, let us consider the English verb *announce*, which is relatively flexible regarding its subcategorization behavior; Boas (2009) analyzed instances of the verb and found

| FEE | Annotated German frames | Count | Most similar English frames |
|-------------------|-------------------------|-------|--------------------------------------|
| <i>ankündigen</i> | HERALDING-DE | 85 | STATEMENT-EN , REPARATION-EN, |
| | OMEN-DE | 1 | OMEN-EN, REST-EN |
| | EVIDENCE-DE | 1 | |

Table 9: SALSA FEEs and frames for translations of the English STATEMENT FEE *announce* (left), and most similar English frame embeddings (right)

that the STATEMENT-EN frame, which can express the SPEAKER, MEDIUM, and MESSAGE roles, is able to cover all of the following uses (Boas, 2009):

- (1) a. [The CEO SPEAKER] *announced* [that the company would be acquired MESSAGE].
- b. [The press report MEDIUM] *announced* [that the company would be acquired MESSAGE].
- c. [The CEO SPEAKER] *announced* [that the company would be acquired MESSAGE] [by email MEDIUM].

The closest German translation of *announce* is *ankündigen*, which can likewise express each of the alternations above. Consequently, we would expect the instances of *ankündigen* in SALSA to be annotated with STATEMENT-DE.

However, as the left-hand cell of Table 9 shows, SALSA decided to annotate the overwhelming majority of the *ankündigen* instances with the frame HERALDING. HERALDING is a specific frame involving a COMMUNICATOR informing the public about a future course of action (EVENT), and in English HERALDING can only be evoked by the verb *herald*. Presumably, the reason for the German annotation is that the very typical use of *ankündigen* with a direct object NP expresses not just a statement event but a commitment by the speaker to carry out the action described by the NP:

- (2) [Regierung in Rom COMMUNICATOR] *kündigt* [Preisstopp und Sparprogramm EVENT] *an*. (s1494)
(Government in Rome announces price stop and austerity program)

Unfortunately, in FrameNet, there is no direct connection of any type between STATEMENT and HERALDING. What our cross-lingual embeddings can provide in this case is an insight into the relationship between these two frames. We do this by retrieving the most similar English frame embeddings for the SALSA *ankündigen* embedding. The result that we find is the most similar frame to *ankündigen* is STATEMENT-EN. This indicates that the usage of *ankündigen* in the SALSA corpus is, even if not labeled as such, semantically still very similar to STATEMENT-EN. Arguably, this pattern - a predicate and its direct translations that belong to two different frames, provides evidence that the involved frames share some kind of conceptual relationship (Sikos and Padó, 2018). In other words, the cross-lingual model has identified a potential gap in the FrameNet frame relations.

7 Conclusion

In this paper, we have looked at the question of the cross-lingual applicability of FrameNet semantic frames. Our starting hypothesis was that multilingual vector spaces can serve as a rich source of information on this topic. A frame vector space can represent frame-evoking elements and frames as vectors, which are readily comparable by geometric distances, thus providing similarity judgments between them.

In our experiments, we have constructed a joint, bilingual space by applying word embedding methods to the annotations of the English FrameNet corpus and the German SALSA corpus. The outcome of three specific analyses that we have carried out – specificity, frame similarity, and translational variation – provides overall positive evidence for this hypothesis: In many cases, our findings correspond well to previous linguistic findings reported, for example, in Boas (2009). Based on our results, we believe that vector space models can be used to help the linguistic/lexicographic work on cross-lingual frame semantics by guiding lexicographers towards potential frame mismatches and cross-lingual applicability.

That said, the picture that emerged is somewhat more nuanced than we anticipated. As usual in corpus-based research, one must be aware that the similarity structure of the embeddings is not always directly interpretable in terms of frame generalizability. Among the factors that impact embeddings are “usual

suspects” like frequency (low-frequency embeddings are potentially noisy, cf. Table 9) and polysemy (the sense distribution of FEEs influences vectors), but also the annotation decisions of individual annotation projects (cf. SALSA’s incomplete coverage of frames in terms of FEEs, and the decision to annotate *ankündigen* with HERALDING-DE). A deeper investigation of these factors is a topic for future work.

Another clear avenue for future investigation is to apply our analysis to compare semantic annotations in multiple languages, which is fast becoming a possibility as new, multilingual FrameNets emerge.

References

- Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of ACL/COLING*, pages 86–90, Montreal, QC.
- Marco Baroni and Alessandro Lenci. 2010. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36(4):673–721.
- Marco Baroni, Georgiana Dinu, and Germán Kruszewski. 2014. Don’t count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of ACL*, pages 238–247, Baltimore, MD.
- Daniel Bauer, Hagen Fürstenu, and Owen Rambow. 2012. The dependency-parsed FrameNet corpus. In *Proceedings of LREC*, pages 3861–3867.
- Hans C Boas. 2005. Semantic frames as interlingual representations for multilingual lexical databases. *International Journal of Lexicography*, 18(4):445–478.
- Hans C Boas. 2009. *Multilingual FrameNets in computational lexicography – Methods and applications*. De Gruyter.
- Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. 2009. Using FrameNet for the semantic analysis of German: Annotation, representation, and automation. In Hans C. Boas, editor, *Multilingual FrameNets in Computational Lexicography – Methods and Applications*, pages 209–244. De Gruyter.
- Alexis Conneau, Guillaume Lample, Marc’Aurelio Ranzato, Ludovic Denoyer, and Hervé Jégou. 2017. Word translation without parallel data. *arXiv:1710.04087*.
- Michael Ellsworth, Katrin Erk, Paul Kingsbury, and Sebastian Padó. 2004. PropBank, SALSA and FrameNet: How design determines product. In *Proceedings of the LREC workshop on Building Lexical Resources From Semantically Annotated Corpora*, pages 17–23, Lisbon, Portugal.
- Katrin Erk, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. 2003. Towards a resource for lexical semantics: A large German corpus with extensive semantic annotation. In *Proceedings of ACL*, pages 537–544, Sapporo, Japan.
- Charles J Fillmore, Christopher R Johnson, and Miriam R L Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.
- Zellig S Harris. 1954. Distributional structure. *Word*, 10(2-3):146–162.
- Karl Moritz Hermann, Dipanjan Das, Jason Weston, and Kuzman Ganchev. 2014. Semantic frame identification with distributed word representations. In *Proceedings of ACL*, pages 1448–1458, Baltimore, MD.
- Birte Lönneker-Rodman. 2007. Multilinguality and FrameNet. Technical report, International Computer Science Institute.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. *arXiv:1301.3781*.
- Tomas Mikolov, Quoc V Le, and Ilya Sutskever. 2013b. Exploiting similarities among languages for machine translation. *arXiv:1309.4168*.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013c. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

- Sebastian Padó. 2007. Translational equivalence and cross-lingual parallelism: The case of FrameNet frames. In *Proceedings of the NODALIDA workshop on building Frame semantics resources for Scandinavian and Baltic languages*, pages 39–46, Tartu, Estonia.
- Sebastian Ruder, Ivan Vulić, and Anders Søgaard. 2017. A survey of cross-lingual word embedding models. *arXiv:1706.04902*.
- Jennifer Sikos and Sebastian Padó. 2018. Framenet’s ‘using’ relation as source of concept-driven paraphrases. *Constructions and Frames*, 10(1). In press.
- Carlos Subirats. 2009. Spanish FrameNet: A frame semantic analysis of the Spanish lexicon. In Hans C. Boas, editor, *Multilingual FrameNets in Computational Lexicography – Methods and Applications*, pages 135–162. De Gruyter.
- Sara Tonelli and Emanuele Pianta. 2008. Frame information transfer from English to Italian. In *Proceedings of LREC*, pages 2252–2256, Marrakech, Morocco.
- Tiago Timponi Torrent, Lars Borin, and Collin Baker, editors. 2018. *Proceedings of the LREC 2018 International FrameNet Workshop on Multilingual Framenets and Constructicons*, Miyazaki, Japan.
- Peter D Turney and Patrick Pantel. 2010. From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research*, 37(1):141–188.
- Shyam Upadhyay, Manaal Faruqui, Chris Dyer, and Dan Roth. 2016. Cross-lingual models of word embeddings: An empirical comparison. In *Proceedings of ACL*, pages 1661–1670, Berlin, Germany.
- Piek Vossen, Antske Fokkens, Isa Maks, and Chantal van Son. 2018. Towards an open Dutch FrameNet lexicon and corpus. In *Proceedings of the LREC 2018 Workshop International FrameNet Workshop 2018 : Multilingual Framenets and Constructicons*, pages 75–80, Miyazaki, Japan.